

10423.204-WO.ST25.txt
SEQUENCE LISTING

<110> Lassen, Soren Flensted
<120> Improved proteases and methods for producing them
<130> 10423.204-WO-DK
<160> 53
<170> PatentIn version 3.2
<210> 1
<211> 1062
<212> DNA
<213> Nocardiopsis sp. NRRL 18262

<220>
<221> misc_feature
<222> (1)..(495)
<223> Encodes the pro-region shown in positions -165 to -1 of SEQ ID NO:43.

<220>
<221> misc_feature
<222> (496)..(1059)
<223> Encodes the mature region shown in positions 1-188 of SEQ ID NO:43.

<400> 1
gctactggag cattacctca gtctcctaca cctgaagcag atgcagtatc gatgcaagaa 60
gcattacaac gtgatcttga tcttacatca gctgaagctg aggaattact tgctgcacaa 120
gatacagcct ttgaagttga tgaagctgcc gctgaagcag ctggtgatgc atatgggtg 180
tcagttatcg atactgaatc actcgaactt actgtactag tgaccgatgc agcagctgtt 240
gaagctgttg aagccacagg tgcaggtaca gagctcgtat cttatggtat tggatggatta 300
gatgagatcg tacaagagct taatgcagct gatgccgttc caggtgttagt tggatggtat 360
cctgatgttag caggtgatac tttgtctta gaagttctt aaggctctgg agctgtatgtt 420
tctggacttt tagcagacgc aggagtcgat gcatccgcgg ttgaagtgtac cacgtcagat 480
cagccctgaac tctatgccga tatcatttggaa ggcctagcgt acacaatggg tggtcgctgc 540
agcgttaggat ttgcagccac aaatgcagct ggacaacctg gcttcgtgac agctggacat 600
tgcggccgcg tcggtagaca gtttactatc ggcaatggaa gaggtgtctt tgagcaaagc 660
gtatcccgg ggaatgtatgc tgccttcgtt agaggtacgt ccaactttac gcttactaac 720
ttatgtatcta gatacaacac tggcggatat gcaactgttag caggtcacaa tcaagcacct 780
attggctcta gcgctgtccg ctcagggtcg actacaggat ggcattgtgg aaccattcaa 840
gctagaggc agagcgtgag ctatcctgaa ggtaccgtaa cgaacatgac tcgtacgact 900
gtatgtgcag aaccagggtga ctctggaggt tcatatatca gcggtacgca agcgcaaggc 960
gttacacctcg gtggatccgg taactgttagg acaggtggca caacgttcta ccaggaagtg 1020
acaccgatgg tgaactcttgg gggagttaga ctccgtacat aa 1062

10423.204-WO.ST25.txt

<210> 2
 <211> 1143
 <212> DNA
 <213> Artificial sequence

 <220>
 <223> A synthetic 10R gene (10Rsynt-15) encoding a S2A protease denoted "10R" fused by PCR in frame to the signal peptide encoding sequence of a heterologous protease, Savinase.

 <400> 2
 atgaagaaac cgttggggaa aattgtcgca agcaccgcac tactcatttc tgttgctttt 60
 agttcatcga tcgcacatggc tgctactgga gcattacctc agtctcctac acctgaagca 120
 gatgcagtat cgatgcaaga agcattacaa cgtgatcttg atcttacatc agctgaagct 180
 gaggaattac ttgctgcaca agatacagcc tttgaagttt atgaagctgc cgctgaagca 240
 gctggtgatg catatggtgg ttcaagtattt gatactgaat cactcgaact tactgtacta 300
 gtgaccgatg cagcagctgt tgaagctgtt gaagccacag gtgcaggtac agagctcgta 360
 tcttatggta ttgatggatt agatgagatc gtacaagagc ttaatgcagc tgatgccgtt 420
 ccaggtgttag ttggatggta tcctgatgta gcaggtgata ctgttgtctt agaagttctt 480
 gaaggctctg gagctgatgt ttctggactt ttagcagacg caggagtgcga tgcatccgcg 540
 gttgaagtga ccacgtcaga tcagcctgaa ctctatgccg atatcattgg aggccctagcg 600
 tacacaatgg gtggtcgtcg cagcgttagga tttgcagcca caaatgcagc tggacaacct 660
 ggcttcgtga cagctggaca ttgcggccgc gtcggcacac aggttactat cggcaatgg 720
 agaggtgtct ttgagcaaag cgtatccc gggaatgatg ctgccttcgt tagaggtacg 780
 tccaacttta cgcttactaa cttagtatct agatacaaca ctggcggata tgcaactgta 840
 gcaggtcaca atcaagcacc tattggctct agcgtctgcc gctcagggtc gactacagga 900
 tggcattgtg gaaccattca agctagaggt cagagcgtga gctatcctga aggtaccgta 960
 acgaacatga ctcgtacgac tgtatgtgca gaaccaggtg actctggagg ttcatatatc 1020
 agcggtacgc aagcgcaagg cgttacactca ggtggatccg gtaactgttag gacaggtggc 1080
 acaacgttct accaggaagt gacaccgatg gtgaactctt ggggagttag actccgtaca 1140
 taa
 1143

<210> 3
 <211> 8
 <212> PRT
 <213> Artificial sequence

 <220>
 <223> C-terminal amino acid tail expressed as fusion to protease of the invention.

<400> 3

Gln Ser His Val Gln Ser Ala Pro
1 5

<210> 4

10423.204-WO.ST25.txt

<211> 24
<212> DNA
<213> Artificial sequence

<220>
<223> Polynucleotide encoding a C-terminal amino acid tail expressed as fusion to protease of the invention.

<400> 4
caatcgcatg ttcaatccgc tcca 24

<210> 5
<211> 4
<212> PRT
<213> Artificial sequence

<220>
<223> C-terminal amino acid tail expressed as fusion to protease of the invention.

<400> 5

Gln Ser Ala Pro
1

<210> 6
<211> 12
<212> DNA
<213> Artificial sequence

<220>
<223> Polynucleotide encoding a C-terminal amino acid tail expressed as fusion to protease of the invention.

<400> 6
caatcggttc ct 12

<210> 7
<211> 2
<212> PRT
<213> Artificial sequence

<220>
<223> C-terminal amino acid tail expressed as fusion to protease of the invention.

<400> 7

Gln Pro
1

<210> 8
<211> 6
<212> DNA
<213> Artificial sequence

<220>
<223> Polynucleotide encoding a C-terminal amino acid tail expressed as fusion to protease of the invention.

<400> 8
caaccca 6

10423.204-WO.ST25.txt

<210> 9
<211> 1
<212> PRT
<213> Artificial sequence

<220>
<223> C-terminal amino acid tail expressed as fusion to protease of the invention.

<400> 9

Pro
1

<210> 10
<211> 3
<212> DNA
<213> Artificial sequence

<220>
<223> Polynucleotide encoding a C-terminal amino acid tail expressed as fusion to protease of the invention.

<400> 10
cca 3

<210> 11
<211> 45
<212> DNA
<213> Artificial sequence

<220>
<223> Primer #252639

<400> 11
catgtgcatg tgggtaccgc aacgttcgca gatgctgctg aagag 45

<210> 12
<211> 44
<212> DNA
<213> Artificial sequence

<220>
<223> Primer #251992

<400> 12
catgtgcatg tggtcgaccg attatggagc ggattgaaca tgcg 44

<210> 13
<211> 44
<212> DNA
<213> Artificial sequence

<220>
<223> Primer #179541

<400> 13
gcgttgagac gcgcggccgc gagcgccgtt tggctgaatg atac 44

<210> 14
<211> 43
<212> DNA
<213> Artificial sequence

10423.204-WO.ST25.txt

<220>
<223> Primer #179542

<400> 14
gcgtttagac agctcgagca gggaaaaatg gaaccgcttt ttc 43

<210> 15
<211> 64
<212> DNA
<213> Artificial sequence

<220>
<223> Primer #179539

<400> 15
ccattttagtc agaatttactt ggcgcgtt ttacaaccat tgcggaaaat agtcataaggc 60
atcc 64

<210> 16
<211> 60
<212> DNA
<213> Artificial sequence

<220>
<223> Primer #179540

<400> 16
ggatccagat ctggtacccg ggtcttagagt cgacgcggcg gttcgcgtcc ggacagcaca 60

<210> 17
<211> 37
<212> DNA
<213> Artificial sequence

<220>
<223> Primer #179154

<400> 17
gttgtaaaac gacggccagt gaattctgat caaatgg 37

<210> 18
<211> 37
<212> DNA
<213> Artificial sequence

<220>
<223> Primer #179153

<400> 18
ccgcgtcgac actagacacg ggtacctgat ctagatc 37

<210> 19
<211> 22
<212> DNA
<213> Artificial sequence

<220>
<223> Primer #317

<400> 19
tggcgtcaatc ggtaccatgg gg 22

10423.204-WO.ST25.txt

<210> 20
<211> 40
<212> DNA
<213> Artificial sequence

<220>
<223> Primer #139 NotI

<400> 20
catgtgcatt cggccgcatt aacgcgttgc cgcttctgcg 40

<210> 21
<211> 7443
<212> DNA
<213> Artificial sequence

<220>
<223> Sequence of plasmid pMB1508

<400> 21
tcgcgcgtt cggtgatgac ggtaaaaacc tctgacacat gcagctcccg gagacggtca 60
cagcttgtct gtaagcggat gccgggagca gacaagcccg tcagggcgcg tcagcgggtg 120
ttggcgggtg tcggggctgg cttaactatg cggcatcaga gcagattgtt ctgagagtgc 180
accatatgcg gtgtgaaata ccgcacagat gcgtaaggag aaaataccgc atcaggcgcc 240
attcgccatt caggctgcgc aactgttggg aagggcgatc ggtgcgggccc tcttcgctat 300
tacccagct ggcgaaaggg ggtatgtctg caaggcgatt aagttggta acgcagggt 360
tttccagtc acgacgttgtt aaaaacgacgg ccagtgaatt cgataaaaagt gcttttttg 420
ttgcaattga agaattatta atgttaagct taattaaaga taatatctt gaattgtAAC 480
gccccctcaaa agtaagaact acaaaaaaaag aatacgttat atagaaatat gtttgaacct 540
tcttcagatt acaaataata tcggacggac tctacctcaa atgcttatct aactatagaa 600
tgacatacaa gcacaaccctt gaaaatttga aaatataact accaatgaac ttgttcatgt 660
gaattatcgc tgtatTTAAT tttctcaatt caatataaa tatGCCAATA cattgttaca 720
agtagaaatt aagacaccct tgatagcctt actataccta acatgatgtt gtattaaatg 780
aatatgtaaa tatattttagt ataagaagcg acttattttt aatcattaca tattttctt 840
ttggaatgtat taagattcca atagaatagt gtataaatta tttatcttga aaggaggat 900
gcctaaaaac gaagaacatt aaaaacatat atttgcaccg tctaattggat ttatgaaaaa 960
tcattttatc agtttggaaaa ttatgttata tggagctctg aaaaaaagga gaggataaag 1020
aatgaagaaa ccgttggga aaattgtcgc aagcaccgca ctactcattt ctgttgcTTT 1080
tagttcatcg atcgcatcgg ctgctgaaga agcaaaagaa aaatattaa ttggctttaa 1140
tgagcaggaa gctgtcagtg agtttgtaga acaagttagag gcaaattgacg aggtcgccat 1200
tctctctgag gaagaggaag tcgaaattga attgcttcat gaatttgaaa cgattcctgt 1260
tttatccgtt gatgttaagcc cagaagatgt ggacgcgcTTT gaactcgatc cagcgatttc 1320
ttatattgaa gaggatgcag aagtaacgac aatggcgcaaa tcggtaccat ggggtatatc 1380

10423.204-WO.ST25.txt

aacgcgttaa	tccgcggata	tatagcggcc	gcagatctgg	gaccaataat	aatgactaga	1440
gaagaaaagaa	tgaagattgt	tcatgaaatt	aaggaacgaa	tattggataa	agtgggatat	1500
ttttaaaata	tatatttatg	ttacagtaat	attgactttt	aaaaaaaggat	tgattctaatt	1560
gaagaaaagca	gacaagtaag	cctcctaaat	tcactttaga	taaaaattta	ggaggcatat	1620
caaataact	ttaataaaaat	tgattnagac	aattggaaga	aaaaaagagat	atttaatcat	1680
tattnaacc	aacaaacgac	ttttagtata	accacagaaa	ttgatattag	tgttttatac	1740
cgaaacataa	aacaagaagg	atataaattt	taccctgcatt	ttattnctt	agtgacaagg	1800
gtgataaaact	caaatacagc	ttttagaact	ggttacaata	gcgacggaga	gttaggttat	1860
tggataagt	taggccact	ttatacaatt	tttgatggtg	tatctaaaac	attctctgg	1920
attnggactc	ctgtaaagaa	tgacttcaaa	gagttttatg	atttatacct	ttctgtatgta	1980
gagaaatata	atggttcggg	gaaattgttt	cccaaaacac	ctatacctga	aaatgctttt	2040
tcttttcta	ttattccatg	gacttcattt	actgggttt	actttaatata	caataataat	2100
agtaattacc	ttctaccat	tattacagca	ggaaaattca	ttaataaagg	taattcaata	2160
tattnaccgc	tatcttaca	ggtacatcat	tctgtttgt	atggttatca	tgcaggattg	2220
tttatgaact	ctattcagga	attgtcagat	aggcctaatt	actggctttt	ataatatgag	2280
ataatgccga	ctgtactttt	tacagtcggt	tttctaacga	tacattaata	ggtacgaaaa	2340
agcaactttt	tttgcgctta	aaaccagtca	taccaataac	ttaagggtaa	ctagcctcgc	2400
cggaaagagc	gaaaatgcct	cacatttgcg	ccacctaataa	aggagcgtt	tacatatgag	2460
ttatgcagtt	tgttagaatgc	aaaaagtgaa	atcagctgga	ctaaaagggg	ccgcagagta	2520
gaatggaaaa	ggggatcgg	aaacaagtat	ataggaggag	acctatttt	ggcttcagaa	2580
aaagacgcag	gaaaacagtc	agcagtaaaag	cttgcattt	tgcttattac	tgtcgctgt	2640
ggactaatca	tctggtttat	tccgcctccg	tccggacttg	aacctaatt	ttggcatttg	2700
tttgcgattt	ttgtcgcaac	aattatcggc	tttatctcca	agcccttgc	aatgggtgca	2760
attgcaattt	ttgcatttgc	ggttactgca	ctaactggaa	cactatcaat	tgaggataca	2820
ttaagcggat	tcgggaataa	gaccatttgg	cttacgttca	tcgcatttctt	tatttcccgg	2880
ggatttatca	aaaccggct	cggcgcgaga	atttcgtatg	tattcgttca	gaaattcgg	2940
aaaaaaaaacc	ttggactttc	ttattcactg	ctattcagt	attnaatact	ttcacctgct	3000
attccaagta	atacggcgcg	tgcaggaggc	attatatttc	ctattatcag	atcattatcc	3060
gaaacattcg	gatcaagccc	ggcaaattgga	acagagagaa	aaatcggtgc	attcttatta	3120
aaaaccggtt	ttcaggggaa	tctgatcaca	tctgtatgt	tcctgacagc	gatggcggcg	3180
aacccgctga	ttgccaagct	ggcccatgat	gtcgacgggg	tggacttaac	atggacaagc	3240
tggcaattt	ccgcatttgt	accgggactt	gtaagcttaa	tcatcagcc	gcttgcgtt	3300
tacaaactgt	atccgcccga	aatcaaagaa	acaccggatg	cggcggaaat	cgcaacagaa	3360
aaactgaaag	aatgggacc	gttcaaaaaa	tcggagctt	ccatggttat	cgtgtttctt	3420

10423.204-WO.ST25.txt

tttgtgcttg	tgctgtggat	tttggcggc	agcttcaaca	tcgacgctac	cacaaccgca	3480								
ttgatcggtt	tggccgttct	cttattatca	caagttctga	cttgggatga	tatcaagaaa	3540								
gaacagggcg	cttgggatac	gctcacttgg	tttgcggcgc	ttgtcatgct	cgc当地acttc	3600								
ttgaatgaat	taggcattgtt	gtcttggttc	agtaatgcca	tgaaatcatc	cgtatcaggg	3660								
ttctcttgg	ttgtggcatt	catcatttttta	attgtgtgt	attattactc	tcactatattc	3720								
tttgc当地gt	cgacagccca	catcagtgcg	atgtattcag	catttttggc	tgtcgtcgtg	3780								
gcagcggg	gag	caccgc当地ct	tttagcagcg	ctgaggctcg	cgttcatcag	3840								
gggtcaacga	ctca	ctcactacgg	ttctggagcg	gctccgtct	tcttcggagc	3900								
ccgcaaggca	aatgg	gttgc当地gtc	catcgattt	atcctgtcga	ttgttcatat	3960								
cttgc当地gt	gcgg	attat	gtggaaagta	ctaggaat	at gtagaaaga	aaaaggcaga	4020							
cgc当地gtc	ctt	tttactcct	tcgt	taagaaa	atggat	tttggaaa	4080							
aaattccctg	tgaaaatgg	tatgatctag	gtagaaagga	cggctgg	tg	tggtgaaa	4140							
aagcgg	ttcc	attttccct	gcaa	acaaaa	ataatgg	tgattgcggc	4200							
tttgc当地tt	gtgt	gtcgtac	cattacgtt	gccgtt	cagc	atcacaggg	4260							
caggcag	agct	ggcggt	tcaa	acggcg	agaacc	attt	ccat	cctat	atg	gccgtt	aaa	4320		
gagctcatt	tg	agagaaaga	cgg	acatgcg	gct	cagacgc	aagagg	tc	at	gat	tga	acaaat	4380	
aaagaacaga	ctgg	tg	ccattt	at	ttt	gaac	gg	aa	gg	ttt	gg	gt	aaaaggaga	4440
gcctctggaa	aa	agcgg	attt	aa	agaa	actg	gagc	gcag	ca	gaa	at	ttt	ggcggt	4500
tcgcatgtt	ctgaa	acaaa	agcgg	atg	ga	cga	agagt	ga	tc	agagg	gg	cg	cgccg	4560
ataaaaagaac	aga	agggata	cag	cca	agt	atc	ggc	agc	tg	tct	gtt	ga	tttctgca	4620
acggagacag	agc	aaagcat	caa	aaag	cat	tt	gagaa	attt	tg	ag	tg	tg	gt	4680
gtactgctgc	tcgg	attt	t	ggc	ccg	cc	tg	ct	gg	cg	ca	tt	tg	4740
ctcggg	ctt	g	a	ccg	cat	g	at	gc	gg	ct	at	tc	tg	4800
gcgattc	gag	ggatt	t	g	cc	ac	ca	at	g	tc	cc	tt	at	4860
tcggcggcc	agat	gct	g	cc	gag	c	c	t	gt	gt	at	tt	cc	4920
atgccgg	gg	agg	ct	g	ct	g	at	cc	tt	at	tt	cc	tt	4980
gta	acgt	gtca	ac	gt	ca	gt	tt	tt	tt	tt	tt	tt	tt	5040
gcgtat	gg	gg	tt	gt	gt	ca	tt	tt	tt	tt	tt	tt	tt	5100
ttgac	ag	agg	tt	cg	ca	aa	tt	tt	tt	tt	tt	tt	tt	5160
aagctt	tt	at	g	tt	tt	tt	tt	tt	tt	tt	tt	tt	tt	5220
gg	tc	at	g	tc	at	g	tc	at	g	tc	at	g	tc	5280
ccgg	aa	gt	ta	aa	g	cc	tt	gg	tt	gg	tt	gg	tt	5340
cg	tt	gc	g	ct	cc	cc	ct	cc	at	cc	at	cc	at	5400
tcggcca	ac	g	cg	gg	gg	g	gc	gg	tt	tg	tt	tg	tt	5460

10423.204-WO.ST25.txt

ctgactcgct	gcgctcggtc	gttcggctgc	ggcgagcggt	atcagctcac	tcaaaggcgg	5520
taatacggtt	atccacagaa	tcagggata	acgcaggaaa	gaacatgtga	gcaaaaggcc	5580
agcaaaaggc	caggaaccgt	aaaaaggccg	cgttgctggc	gttttccat	aggctccgccc	5640
cccctgacga	gcatcacaaa	aatcgacgct	caagtcagag	gtggcgaaac	ccgacaggac	5700
tataaagata	ccagggcgaaa	ccccctggaa	gctccctcg	gctcgtctcct	gttccgacccc	5760
tgcccgttac	cggataacctg	tccgcctttc	tcccttcggg	aagcgtggcg	ctttctcata	5820
gctcacgctg	taggtatctc	agttcgggtgt	aggtcgttcg	ctccaagctg	ggctgtgtgc	5880
acgaaccccc	cgttcagccc	gaccgctgctg	ccttatccgg	taactatcg	cttgagtcac	5940
acccggtaag	acacgactta	tcgcccactgg	cagcagccac	tggtaacagg	attagcagag	6000
cgaggtatgt	aggcggtgct	acagagttct	tgaagtggtg	gcctaactac	ggctacacta	6060
gaaggacagt	atttggtatac	tgcgctctgc	tgaagccagt	taccttcgga	aaaagagttg	6120
gtagctttg	atccggcaaa	caaaccaccg	ctggtagcgg	tggtttttt	gttgcaagc	6180
agcagattac	gcgcagaaaa	aaaggatctc	aagaagatcc	tttgcatttt	tctacgggg	6240
ctgacgctca	gtggaacgaa	aactcacg	aagggatttt	ggtcatgaga	ttatcaaaaa	6300
ggatcttcac	ctagatcctt	ttaaattaaa	aatgaagttt	taaatcaatc	taaagtat	6360
atgagtaaac	ttggctgac	agttaccaat	gcttaatcag	tgaggcacct	atctcagcga	6420
tctgtctatt	tcggtcatcc	atagttgcct	gactccccgt	cgtgtagata	actacgatac	6480
gggagggctt	accatctggc	cccagtgctg	caatgatacc	gcgagaccca	cgctcaccgg	6540
cacccggattt	atcagcaata	aaccagccag	ccggaagggc	cgagcgcaga	agtggcctg	6600
caactttatc	cgccctccatc	cagtctatta	attgttgcgg	ggaagctaga	gtaagtagtt	6660
cgccagttaa	tagttgcgc	aacgttgc	ccattgctac	aggcatcgtg	gtgtcacgct	6720
cgtcgttgg	tatggcttca	ttcagctccg	gttcccaacg	atcaaggcga	gttacatgt	6780
ccccccatgtt	gtgcaaaaaaa	gcggtagct	ccttcgtcc	tccgatcg	gtcagaagta	6840
agttggccgc	agtgttatca	ctcatggta	tggcagcact	gcataattct	cttactgtca	6900
tgcccatccgt	aagatgcttt	tctgtgactg	gtgagtagtc	aaccaagtca	ttctgagaat	6960
agtgtatgcg	gchgaccgagt	tgctctgccc	cgccgtcaat	acgggataat	accgcgccac	7020
atagcagaac	tttaaaagtg	ctcatcattg	gaaaacgttc	ttcggggcga	aaactctcaa	7080
ggatcttacc	gctgttgc	tccagttcga	tgtaacccac	tcgtgcaccc	aactgatctt	7140
cagcatcttt	tacccatcacc	agcgtttctg	ggtgagcaaa	aacaggaagg	caaaatgccc	7200
caaaaaagg	aataagggcg	acacggaaat	gttgaatact	catactcttc	cttttcaat	7260
attattgaag	catttatcag	gttattgtc	tcatgagcgg	atacatattt	aatgtattt	7320
agaaaaataa	acaaataggg	gttccgcgc	catttccccg	aaaagtgcac	cctgacgtct	7380
aagaaaccat	tattatcatg	acattaacct	ataaaaatag	gctgtatcag	aggccctttc	7440
gtc						7443

10423.204-WO.ST25.txt

<210> 22
 <211> 5718
 <212> DNA
 <213> Artificial sequence

 <220>
 <223> Sequence of MB1510 genomic integration region

 <400> 22

gagcgcgtt	tggctgaatg	ataacaacagt	ctcacttcct	tactgcgtct	ggttgc	aaaaa	60
acgaagaagc	aaggattccc	ctcgcttctc	atttgccta	tttattatac	actttttt	aa	120
gcacatctt	ggcgcttgtt	tcactagact	tgatgcctct	aatcttg	tc	caagtgtcac	180
ggtccgcac	atagacttgt	ccat	tttca	ccgctt	gag	attttccag	240
ttttccactc	atctacaatg	gtttgc	ctt	cg	ttgctga	gatgaac	300
cgat	tttgc	caattg	ctca	agg	ctgac	tttgc	360
gtgt	aaagcc	tagcattt	ta	aagattt	ctc	cgtcat	420
aggaatccgc	tcttg	caac	cg	cg	ttgttgc	tttgc	480
cttttagatc	gtt	gatgact	ttttt	gt	cg	caagct	540
tat	taatgc	tttag	caatg	tcgt	atcg	tttc	600
ggctt	ttttaa	ttcaat	cg	ggat	tttca	gat	660
gctcagcg	tc	agcgatg	ttt	tttgc	tttgc	tttgc	720
cgctgcgt	gc	ctacagat	gt	taat	ca	tttgc	780
ctttttgtt	gt	cat	ct	gc	at	cccac	840
agaatgaa	ag	ct	caag	caca	cc	accac	900
cttcgtcat	gt	tc	act	ct	gt	tttgc	960
ggctt	gt	at	cc	tttgc	tttgc	tttgc	1020
tcaact	ttt	tttgc	tttgc	tttgc	tttgc	tttgc	1080
tcattatc	at	gt	taa	act	at	at	1140
gaaagct	ttt	tttgc	tttgc	tttgc	tttgc	tttgc	1200
cagcgttacc	gt	gt	tttgc	gac	aaaac	ac	1260
tatc	ac	gt	tttgc	tttgc	tttgc	tttgc	1320
catctgtt	tg	caaa	tttgc	tttgc	tttgc	tttgc	1380
tatcgat	at	gt	tttgc	tttgc	tttgc	tttgc	1440
cggcac	at	gg	tttgc	tttgc	tttgc	tttgc	1500
gttct	tttgc	aa	gtt	tttgc	tttgc	tttgc	1560
tcat	tttgc	tttgc	tttgc	tttgc	tttgc	tttgc	1620
cccgtt	tttgc	tttgc	tttgc	tttgc	tttgc	tttgc	1680
ttct	tttgc	tttgc	tttgc	tttgc	tttgc	tttgc	1740

10423.204-WO.ST25.txt

gaaaacatga aggaggaata tcgacatgaa accagttgt a	aaagagtata caaatgacga	1800
acagctcatg aaagatgtag aggaattgca gaaaatgggt	gttgcgaaag aggatgtata	1860
cgtcttagct cacgacgatg acagaacgga acgcctggct	gacaacacga acgccaacac	1920
gatcgagcc aaagaaacag gtttcaagca cgccgtggga	aatatcttca ataaaaaagg	1980
agacgagctc cgcaataaaa ttcacgaaat cggttttct	gaagatgaag ccgctcaatt	2040
tgaaaaacgc ttagatgaag gaaaagtgc tctcttgc	acagataacg aaaaagtgaa	2100
agcttgggca taaagcaagg aaaaaaccaa aaggccaatg	tcggccttt ggttttttg	2160
cggtcttgc ggtgggattt tgcagaatgc cgcaatagga	tagcggaaaca tttcggttc	2220
tgaatgtccc tcaatttgc attatatttt tgtgataaaat	tggaataaaa tctcacaaaaa	2280
tagaaaatgg gggtacatag tggatgaaaa aagtgtatgtt	agctacggct ttgttttttag	2340
gattgactcc agctggcgcg aacgcagctg atttaggcca	ccagacgttg ggatccaatg	2400
atggctgggg cgctactcg accggcacga cagggcgatc	aaaagcatcc tcctcaaatg	2460
tgtataccgt cagcaacaga aaccagcttgc tctcggcatt	agggaaaggaa acgaacacaa	2520
cgccaaaaat catttatatc aagggAACGA ttgacatgaa	cgtggatgac aatctgaagc	2580
cgcttggcct aaatgactat aaagatccgg agtatgattt	ggacaaatat ttgaaagcct	2640
atgatcctag cacatggggc aaaaaagagc cgctggaaac	acaagaagaa gcgagagcac	2700
gctctcagaa aaaccaaaaa gcacgggtca tggatgat	ccctgcaaac acgacgatcg	2760
tcgggagg gactaacgct aaagtcgtgg gaggaaactt	ccaaatcaag agtataacg	2820
tcattattcg caacattgaa ttccaggatg cctatgacta	ttttccgcaa tgggtgtaaa	2880
acgacggcca gtgaattctg atcaaattgt tcagttagag	cgaagcgaac acttgatttt	2940
ttaattttct atctttata ggtcatttgc gtatacttat	ttgtcctata aactattttag	3000
cagcataata gatttattga ataggtcatt taagttgagc	atattagagg agaaaaatct	3060
tggagaaaata tttgaagaac ccgagatcta gatcaggtac	cgcaacgttc gcagatgctg	3120
ctgaagagat tattaaaaag ctgaaagcaa aaggctatca	attggtaact gtatctcagc	3180
ttgaagaagt gaagaagcag agaggctatt gaataatga	gtagaaagcg ccatatcgac	3240
gctttcttt tggagaaaaa tatagggaaa atggtacttg	ttaaaaattt ggaatattta	3300
tacaatatca tatgtatcac attgaaagga ggggcctgct	gtccagactg tccgctgtgt	3360
aaaaataagg aataaagggg ggttgcattt attttactga	tatgtataat ataatttgtt	3420
taagaaaaatg gagggccct cgaaacgtaa gatgaaacct	tagataaaag tgctttttt	3480
gttgcattt aagaattt aatgttaagc ttaattaaag	ataatatctt tgaattgtaa	3540
cgcctctcaa aagtaagaac tacaaaaaaa gaatacgta	tatgaaata tgttgaacc	3600
ttcttcagat tacaatatata ttccggacgga ctctacctca	aatgcttatac taactataga	3660
atgacatatac agcacaacct tgaaaattt gaaatataac	taccaatgaa cttgttcatg	3720
tgaattatcg ctgtatttaa tttctcaat tcaatatata	atatgccaat acattgttac	3780

10423.204-WO.ST25.txt

aagttagaaat	taagacaccc	ttgatagcct	tactatacct	aacatgatgt	agtattaaat	3840
gaatatgtaa	atatatttat	gataagaagc	gacttattta	taatcattac	atattttct	3900
attggaatga	ttaagattcc	aatagaatag	tgtataaaatt	atttatcttgc	aaaggaggga	3960
tgcctaaaaa	cgaagaacat	taaaaacata	tatttgcacc	gtctaatttgc	tttatgaaaa	4020
atcattttat	cagtttggaaa	attatgtattt	atggagctct	gaaaaaaaaagg	agaggataaa	4080
gagaaaaggg	gatcgaaaaa	caagtatata	ggaggagacc	tatttatggc	ttcagaaaaa	4140
gacgcaggaa	aacagtccgc	agtaaagctt	gttccattgc	tttattactgt	cgctgtggga	4200
ctaattcatct	ggtttattcc	cgctccgtcc	ggacttgaac	ctaaagcttgc	gcatttgcatt	4260
gcgatttttgc	tcgcaacaat	tatcggtttt	atctccaagc	ccttgcatttgc	gggtgcatttgc	4320
gcaatttttgc	cattggcggt	tactgcacta	actggAACAC	tatcaatttgc	ggatACATTA	4380
agcggattcg	ggaataagac	catttggctt	atcggtatcg	cattttttat	ttcccgggga	4440
tttatcaaaa	ccggtctcg	tgcgagaatt	tcgtatgtat	tcgttcagaa	attcgaaaaa	4500
aaaacccttg	gactttctta	ttcactgcta	ttcagtgatt	taataacttgc	acctgctatt	4560
ccaagtaata	cggcgcgtgc	aggaggcatt	atatttctta	ttatcagatc	attatccgaa	4620
acattcggat	caagccccgc	aatggAACAC	gagagaaaaa	tcgggtgcatt	cttattaaaa	4680
accggttttc	aggggaatct	gatcacatct	gctatgttcc	tgacagcgat	ggcggcgaac	4740
ccgctgatttgc	ccaagctggc	ccatgatgtc	gcaggggtgg	acttaacatg	gacaagctgg	4800
gcaattgccc	cgattgtacc	gggacttgc	agcttaatca	tcacgcgcgt	tgtgattttac	4860
aaactgtatc	cgccggaaat	caaagaaaca	ccggatgcgg	cgaaaaatcg	aacagaaaaa	4920
ctgaaagaaa	tgggaccgtt	caaaaaatcg	gagcttcca	tggttatcg	gtttttttgc	4980
gtgcttgc	tgtggatttt	tggcgccagc	ttcaacatcg	acgctaccac	aaccgcatttgc	5040
atcggtttgg	ccgttctctt	attatcacaa	gttctgactt	gggatgatat	caagaaagaa	5100
cagggcgctt	gggatacgct	cacttggtttgc	gcggcgttgc	tcatgctcgc	caacttcttgc	5160
aatgaatttag	gcatggtgc	ttggttcagt	aatgcatttgc	aatcatccgt	atcagggttc	5220
tcttggatttgc	tggcattcat	catttttatttgc	gttgtgtatt	attactctca	ctatttcttgc	5280
gcaagtgcga	cagcccacat	cagtgcgttgc	tattcagcat	ttttggctgt	cgtcgtggca	5340
gcgggcccac	cggccgttttgc	agcagcgcttgc	agcctcggttgc	tcatcagca	cctgttcggg	5400
tcaacactc	actacggtttgc	tggagcggcttgc	ccggtcttgc	tcggagcagg	ctacatccgc	5460
caaggcaaat	ggtggtccat	cgattttatcg	ctgtcgatttgc	ttcatatcat	cgtatggcttgc	5520
gtgatcggcg	gattatggtg	gaaagtacta	ggaatatggt	agaaaagaaaa	aggcagacgc	5580
ggtctgcctt	tttttatttttgc	caactcatttgc	taagaaaatgc	gattttggaaa	aatgagaaaa	5640
ttccctgtga	aaaatggat	gatcttagtgc	gaaaggacgg	ctgggtgtgt	ggtgaaaaag	5700
cggttccattt	tttcccttgc					5718

10423.204-WO.ST25.txt

<210> 23		
<211> 27		
<212> DNA		
<213> Artificial sequence		
<220>		
<223> Primer 1605		
<400> 23		
gacggccagt gaattcgata aaagtgc	27	
<210> 24		
<211> 42		
<212> DNA		
<213> Artificial sequence		
<220>		
<223> Primer 1606		
<220>		
<221> misc_feature		
<222> (13)..(13)		
<223> n is a, c, g, or t		
<220>		
<221> misc_feature		
<222> (16)..(16)		
<223> n is a, c, g, or t		
<400> 24		
ccagatctct atnktntktgt acggagtctta actccccaaag ag	42	
<210> 25		
<211> 1112		
<212> DNA		
<213> Nocardiopsis dassonvillei DSM 43235		
<400> 25		
gcttttagtt catcgatcgc atcggctgct ccggccccc tcccccaagac ccccggtcgcc	60	
gacgacagcg ccgccagcat gaccgaggcg ctcaagcgcg acctcgaccc cacctcggcc	120	
gaggccgagg agcttctctc ggcgcaggaa gccgccatcg agaccgacgc cgaggccacc	180	
gaggccgcgg gcgaggcccta cggcggctca ctgttcgaca ccgagaccct cgaactcacc	240	
gtgctggta ccgacgcctc cggcgatcgag gcgggtcgagg ccaccggagc ccaggccacc	300	
gtcgtctccc acggcaccga gggcctgacc gaggtcggtgg aggaccta cggcgccgag	360	
gttcccgaga gcgtcctcggtt cttgtaccccg gacgtggaga gcgacaccgt cgtggatcgag	420	
gtgctggagg gctccgacgc cgacgtcgcc gccctgctcg ccgacgcccgg tgtggactcc	480	
tcctcggtcc ggggtggagga ggccgaggag gccccgcagg tctacgcccga catcatcgcc	540	
ggcctggcct actacatggg cggccgctgc tccgtcggt tcgcccgcac caacagcgcc	600	
ggtcagcccg gtttcgtcac cgccggccac tgcggcaccg tcggcaccgg cgtgaccatc	660	
ggcaacggca cccggcacctt ccagaactcg gtcttccccg gcaacgacgc cgccttcgtc	720	
cgcggcacct ccaacttacac cctgaccaac ctggtctcgctc gctacaactc cggcggctac	780	
cagtcggta cccggtaccag ccaggccccg gccggctcggtcc cctgggtcgcc tcggcgtcc	840	

10423.204-WO.ST25.txt

accaccggct ggcactgcgg caccatccag gcccgcAACc agaccgtgcg ctacccgcag	900
ggcacccgtct actcgctcac ccgcaccaac gtgtgcgccc agcccgccgat ctcggcggt	960
tcgttcatct cccgctcgca ggcccaggc gtcacccctccg gcggctccgg caactgctcc	1020
gtcggcggca cgacctacta ccaggaggtc accccgatga tcaactcctg gggtgtcagg	1080
atccggacct aatcgcatgt tcaatccgct cc	1112

<210> 26
<211> 48
<212> DNA
<213> Artificial sequence

<220>
<223> Primer 1423

<400> 26
gcttttagtt catcgatcgc atcggctgct ccggccccc tccccag 48

<210> 27
<211> 45
<212> DNA
<213> Artificial sequence

<220>
<223> Primer 1475

<400> 27
ggagcggatt gaacatgcga ttaggtccgg atcctgacac cccag 45

<210> 28
<211> 354
<212> PRT
<213> Nocardiopsis dassonvillei DSM 43235

<220>
<221> PROPEP
<222> (1)..(166)

<220>
<221> mat_peptide
<222> (167)..(354)

<400> 28

Ala Pro Ala Pro Val Pro Gln Thr Pro Val Ala Asp Asp Ser Ala
-165 -160 -155

Ala Ser Met Thr Glu Ala Leu Lys Arg Asp Leu Asp Leu Thr Ser
-150 -145 -140

Ala Glu Ala Glu Glu Leu Leu Ser Ala Gln Glu Ala Ala Ile Glu
-135 -130 -125

Thr Asp Ala Glu Ala Thr Glu Ala Ala Gly Glu Ala Tyr Gly Gly
-120 -115 -110

10423.204-WO.ST25.txt

Ser Leu Phe Asp Thr Glu Thr Leu Glu Leu Thr Val Leu Val Thr Asp
 -105 -100 -95

Ala Ser Ala Val Glu Ala Val Glu Ala Thr Gly Ala Gln Ala Thr Val
 -90 -85 -80 -75

Val Ser His Gly Thr Glu Gly Leu Thr Glu Val Val Glu Asp Leu Asn
 -70 -65 -60

Gly Ala Glu Val Pro Glu Ser Val Leu Gly Trp Tyr Pro Asp Val Glu
 -55 -50 -45

Ser Asp Thr Val Val Val Glu Val Leu Glu Gly Ser Asp Ala Asp Val
 -40 -35 -30

Ala Ala Leu Leu Ala Asp Ala Gly Val Asp Ser Ser Ser Val Arg Val
 -25 -20 -15

Glu Glu Ala Glu Glu Ala Pro Gln Val Tyr Ala Asp Ile Ile Gly Gly
 -10 -5 -1 1 5

Leu Ala Tyr Tyr Met Gly Gly Arg Cys Ser Val Gly Phe Ala Ala Thr
 10 15 20

Asn Ser Ala Gly Gln Pro Gly Phe Val Thr Ala Gly His Cys Gly Thr
 25 30 35

Val Gly Thr Gly Val Thr Ile Gly Asn Gly Thr Gly Thr Phe Gln Asn
 40 45 50

Ser Val Phe Pro Gly Asn Asp Ala Ala Phe Val Arg Gly Thr Ser Asn
 55 60 65 70

Phe Thr Leu Thr Asn Leu Val Ser Arg Tyr Asn Ser Gly Gly Tyr Gln
 75 80 85

Ser Val Thr Gly Thr Ser Gln Ala Pro Ala Gly Ser Ala Val Cys Arg
 90 95 100

Ser Gly Ser Thr Thr Gly Trp His Cys Gly Thr Ile Gln Ala Arg Asn
 105 110 115

Gln Thr Val Arg Tyr Pro Gln Gly Thr Val Tyr Ser Leu Thr Arg Thr
 120 125 130

Asn Val Cys Ala Glu Pro Gly Asp Ser Gly Gly Ser Phe Ile Ser Gly
 135 140 145 150

Ser Gln Ala Gln Gly Val Thr Ser Gly Gly Ser Gly Asn Cys Ser Val
 155 160 165

10423.204-WO.ST25.txt
 Gly Gly Thr Thr Tyr Tyr Gln Glu Val Thr Pro Met Ile Asn Ser Trp
 170 175 180

Gly Val Arg Ile Arg Thr
 185

<210> 29
 <211> 498
 <212> DNA
 <213> Nocardiopsis dassonvillei DSM 43235
 <400> 29
 gctccggccc cgtccccca gaccccgta gcccacgaca gcgccgccag catgaccgag 60
 gcgctcaagc gcgacctcga cttcacctcg gccgaggccg aggagttct ctcggcgcag
 gaagccgcca tcgagaccga cgccgaggcc accgaggccg cggcgaggc ctacggcgc 120
 tcactgttcg acaccgagac cttcgaaactc accgtgctgg tcaccgacgc ctccgccc 180
 gaggcggtcg aggccaccgg agcccaggcc accgtcgtct cccacggcac cgagggcctg 240
 accgaggtcg tggaggacct caacggcgcc gaggttccc agagcgtcct cggctggta 300
 ccggacgtgg agagcgcac acgtcgtggc gaggtgctgg agggctccga cgccgacgtc 360
 gccgcccctgc tcgcccacgc cggtgtggac tcctcctcgg tccgggtgga ggaggccgag 420
 gaggccccgc aggtctac . 480
 498

<210> 30
 <211> 166
 <212> PRT
 <213> Nocardiopsis dassonvillei DSM 43235
 <400> 30

Ala Pro Ala Pro Val Pro Gln Thr Pro Val Ala Asp Asp Ser Ala Ala
 1 5 10 15

Ser Met Thr Glu Ala Leu Lys Arg Asp Leu Asp Leu Thr Ser Ala Glu
 20 25 30

Ala Glu Glu Leu Leu Ser Ala Gln Glu Ala Ala Ile Glu Thr Asp Ala
 35 40 45

Glu Ala Thr Glu Ala Ala Gly Glu Ala Tyr Gly Gly Ser Leu Phe Asp
 50 55 60

Thr Glu Thr Leu Glu Leu Thr Val Leu Val Thr Asp Ala Ser Ala Val
 65 70 75 80

Glu Ala Val Glu Ala Thr Gly Ala Gln Ala Thr Val Val Ser His Gly
 85 90 95

Thr Glu Gly Leu Thr Glu Val Val Glu Asp Leu Asn Gly Ala Glu Val
 100 105 110

10423.204-WO.ST25.txt
 Pro Glu Ser Val Leu Gly Trp Tyr Pro Asp Val Glu Ser Asp Thr Val
 115 120 125

Val Val Glu Val Leu Glu Gly Ser Asp Ala Asp Val Ala Ala Leu Leu
 130 135 140

Ala Asp Ala Gly Val Asp Ser Ser Ser Val Arg Val Glu Glu Ala Glu
 145 150 155 160

Glu Ala Pro Gln Val Tyr
 165

<210> 31

<211> 1146

<212> DNA

<213> Artificial sequence

<220>

<223> The DNA sequence coding for the pro-region of SEQ ID NO: 29 fused in frame to A1918L2 protease tail-variant encoding gene; whole construct: 10R(proA1918L2).

<400> 31

atgaagaaac	cgttggggaa	aattgtcgca	agcaccgcac	tactcattc	tgttgcttt	60
agttcatcga	tcgcacatcgcc	tgctccggcc	cccggtccccc	agaccccccgt	cgccgacgac	120
agcgccgcca	gcatgaccga	ggcgctcaag	cgcgacacctg	acctcaccc	ggccgaggcc	180
gaggagcttc	tctcggcgca	ggaagccgccc	atcgagaccg	acgcccggagc	caccgaggcc	240
gcggcgagg	cctacggcg	ctcactgttc	gacaccgaga	ccctcgaact	caccgtgctg	300
gtcaccgcac	cctccgcgt	cgaggcggtc	gaggccaccg	gagcccaggc	caccgtcg	360
tcccacggca	ccgagggcct	gaccgaggtc	gtggaggacc	tcaacggcgc	cgaggttccc	420
gagagcgtcc	tcggctggta	cccgacgtg	gagagcaca	ccgtcgttgt	cgaggtgctg	480
gagggctccg	acgcccacgt	cgccgcctg	ctcgccacg	ccggtgtgga	ctccctcctg	540
gtccgggtgg	aggaggccga	ggaggccccg	caggtctatg	ccgatatcat	tggaggccta	600
gcgtacacaa	tgggtggtcg	ctgcagcgt	ggatttgcag	ccacaaatgc	agctggacaa	660
cctggcttcg	tgacagctgg	acattgcggc	cgcgtcggt	cacaggttac	tatcgcaat	720
ggaagaggtg	tcttgagca	aagcgtat	cccggaaatg	atgctgcct	cgtagaggt	780
acgtccaact	ttacgcttac	taacttagta	tctagataca	acactggcgg	atatgcaact	840
gtacgaggc	acaatcaagc	acctattggc	tctagcgtct	gccgctcagg	gtcgactaca	900
ggatggcatt	gtggaaccat	tcaagctaga	ggtcagagcg	tgagctatcc	tgaaggtacc	960
gtaacgaaca	tgactcgtac	gactgtatgt	gcagaaccag	gtgactctgg	aggttcatat	1020
atcagcggta	cgcaagcgca	aggcggtacc	tcaggtggat	ccggtaactg	taggacaggt	1080
ggcacaacgt	tctaccagga	agtgacaccg	atggtaact	cttggggagt	tagactccgt	1140
acataa						1146

10423.204-WO.ST25.txt

<210> 32
 <211> 1068
 <212> DNA
 <213> Nocardiopsis Alba DSM 15647

<400> 32
 gcgaccggcc ccctccccc a g t c c c c a c c g g a a g c c g a g g c c a c c a c c a t g g t c 60
 g a g g c c t c c a g c g c a c c t c g g c t c c c t c t c a g g c c g a c g a g c t c t c g a g g c g 120
 c a g g c c g a g t c c t c g a g a t c g a c g a g g c c g c a c c g c g g c c g a c g c a c c t a c g g c 180
 g g c t c c a t c t c g a c a c c g a c a g c t c a c c t g a c c g t c c t g g t c a c c g a c g c t c c g c c 240
 g t c g a g g c g g t c g a g g c c g c g g c a g g t g g t c t c g c a c c g g a t g g a g g g c 300
 c t g g a g g a g a t c g t c g c g a c c t g a a c g c g g c c g a c g c t c a g c c g g c t g g g c t g g 360
 t a c c c c g a c a t c c g a c a c g t c g t c c t c g a g g t c c t g a g g g c t c g g c c a c 420
 g t g g a c t c c c t g c t c g c g a c c c g g t g g a c a c c g c c g a c g t c a a g g t g g a g a g c a c c 480
 a c c g a g c a g c c c g a g t g a c c g a c a t c a c g g g t c t c g c t a c a c a t g g g t g g g 540
 c g c t g c t c g g t c g g c t t c g c g c g c a c c a a c g c t c c g g c c a g c c g g g t t c g t a c c g c c 600
 g g c c a c t g c g g c a c c g t c g g c a c c c g g t c a c g c a c g g c c a g g g c t c t c g a g 660
 c g t t c c g t c t c c c g g c a a c g a c t c c g c c t c g c a c c g g c a c c g g c t c t c g a g 720
 a c c a a c t c g g t c a c a c c c g g t c a c g c a a c g g c a c c g g c a c c g g c t c t c g a g 780
 g c g g c g a t c g c a g a t c g c g t t c g c g c g a c c c g g c a c c g g c a c c g g c t c t c g a g 840
 g t c c a g g c c c g g c a g a c g g t a c c c a g g g c a c c g g c a c c g g c t c t c g a c c c g c 900
 a c c a a c g t c t c g c c g a g c c g g c a c g g c a c c g g c t c t c g a c c c g c a c c g g c 960
 c a g g g c g t c a c c c g g t g g t c c g g c a a c t g c t c t c g g a c c a c c g g c a c c g g c 1020
 g a g g t c a a c c g a t g c t g a g c a g c t g g g g t c t g a c c c t g c a c c t g c 1068

<210> 33
 <211> 355
 <212> PRT
 <213> Nocardiopsis Alba DSM 15647

<220>
 <221> PROPEP
 <222> (1)..(167)

<220>
 <221> mat_peptide
 <222> (168)..(355)

<400> 33

Ala Thr Gly Pro Leu Pro Gln Ser Pro Thr Pro Asp Glu Ala Glu
 -165 -160 -155

Ala Thr Thr Met Val Glu Ala Leu Gln Arg Asp Leu Gly Leu Ser
 -150 -145 -140

Pro Ser Gln Ala Asp Glu Leu Leu Glu Ala Gln Ala Glu Ser Phe

10423.204-WO.ST25.txt
 -135 -130 -125

Glu Ile Asp Glu Ala Ala Thr Ala Ala Ala Ala Asp Ser Tyr Gly
 -120 -115 -110

Gly Ser Ile Phe Asp Thr Asp Ser Leu Thr Leu Thr Val Leu Val Thr
 -105 -100 -95

Asp Ala Ser Ala Val Glu Ala Val Glu Ala Ala Gly Ala Glu Ala Lys
 -90 -85 -80

Val Val Ser His Gly Met Glu Gly Leu Glu Glu Ile Val Ala Asp Leu
 -75 -70 -65 -60

Asn Ala Ala Asp Ala Gln Pro Gly Val Val Gly Trp Tyr Pro Asp Ile
 -55 -50 -45

His Ser Asp Thr Val Val Leu Glu Val Leu Glu Gly Ser Gly Ala Asp
 -40 -35 -30

Val Asp Ser Leu Leu Ala Asp Ala Gly Val Asp Thr Ala Asp Val Lys
 -25 -20 -15

Val Glu Ser Thr Thr Glu Gln Pro Glu Leu Tyr Ala Asp Ile Ile Gly
 -10 -5 -1 1 5

Gly Leu Ala Tyr Thr Met Gly Gly Arg Cys Ser Val Gly Phe Ala Ala
 10 15 20

Thr Asn Ala Ser Gly Gln Pro Gly Phe Val Thr Ala Gly His Cys Gly
 25 30 35

Thr Val Gly Thr Pro Val Ser Ile Gly Asn Gly Gln Gly Val Phe Glu
 40 45 50

Arg Ser Val Phe Pro Gly Asn Asp Ser Ala Phe Val Arg Gly Thr Ser
 55 60 65

Asn Phe Thr Leu Thr Asn Leu Val Ser Arg Tyr Asn Thr Gly Gly Tyr
 70 75 80 85

Ala Thr Val Ser Gly Ser Ser Gln Ala Ala Ile Gly Ser Gln Ile Cys
 90 95 100

Arg Ser Gly Ser Thr Thr Gly Trp His Cys Gly Thr Val Gln Ala Arg
 105 110 115

Gly Gln Thr Val Ser Tyr Pro Gln Gly Thr Val Gln Asn Leu Thr Arg
 120 125 130

Thr Asn Val Cys Ala Glu Pro Gly Asp Ser Gly Gly Ser Phe Ile Ser
 Page 19

135 140 10423.204-WO.ST25.txt
145

Gly Ser Gln Ala Gln Gly Val Thr Ser Gly Gly Ser Gly Asn Cys Ser
150 155 160 165

Phe Gly Gly Thr Thr Tyr Tyr Gln Glu Val Asn Pro Met Leu Ser Ser
170 175 180

Trp Gly Leu Thr Leu Arg Thr
185

<210> 34
<211> 43
<212> DNA
<213> Artificial sequence

<220>
<223> Primer 1421

<400> 34
gttcatcgat cgcacatcggtc gcgaccggcc ccctccccca gtc

43

<210> 35
<211> 31
<212> DNA
<213> Artificial sequence

<220>
<223> Primer 1604

<400> 35
gcggatccata tcaggtgcgc agggtcagac c

31

<210> 36
<211> 1062
<212> DNA
<213> Nocardiopsis prasina DSM 15648

<400> 36
gccaccggac cgctccca gtcacccacc ccggaggccg acgccgtctc catgcaggag 60
gcgctccagc gcgacctcg cctgaccccg cttgaggccg atgaactgct ggccgcccag 120
gacaccgcct tcgaggtcga cgaggccgcg gccgcggccg ccggggacgc ctacggcggc 180
tccgtcttcg acaccgagac cctggaactg accgtcctgg tcaccgacgc cgcctcggtc 240
gaggctgtgg aggccaccgg cgcggttacc gaactcgctc cctacggcat cgagggcctc 300
gacgagatca tccaggatct caacgcccgc gacgcccgtcc ccggcgtggc cggctggtag 360
ccggacgtgg cgggtgacac cgtcgccctg gaggtcctgg agggttccgg agccgacgtg 420
agcggcctgc tcgcccgcgc cggcggtggac gcctcgcccg tcgaggtgac cagcagtgcg 480
cagcccgagc tctacgcccga catcatcgcc ggtctggcct acaccatggg cggccgtgt 540
tcggtcggat tcgcccggcac caacgcccgc ggtcagcccg gattcgtcac cgcggcgtcac 600
tgtggccgcg tgggcaccca ggtgagcatc ggcaacggcc agggcgttt cgagcagtcc 660
atcttcccg gcaacgacgc cgccttcgtc cgcggcacgt ccaacttcac gctgaccaac 720

10423.204-WO.ST25.txt

ctggtcagcc	gctacaacac	cgccggttac	gccaccgtcg	ccggccacaa	ccagggcccc	780
atcggtcct	ccgtctgccg	ctccggctcc	accaccggct	ggcactgcgg	caccatccag	840
gcccggcc	agtcggtgag	ctaccccgag	ggcaccgtca	ccaacatgac	ccggaccacc	900
gtgtgcggc	agcccggcga	ctccggcggc	tcctacatct	ccggcaacca	ggcccgaggc	960
gtcacctccg	gcggctccgg	caactgcccgc	accggcggga	ccaccttcta	ccaggaggtc	1020
accccatgg	tgaactcctg	ggcggtccgt	ctccggacact	aa		1062

<210> 37
 <211> 353
 <212> PRT
 <213> Nocardiopsis prasina DSM 15648

<220>
 <221> PROPEP
 <222> (1)..(165)

<220>
 <221> mat_peptide
 <222> (166)..(353)

<400> 37

Ala Thr Gly Pro Leu Pro Gln Ser Pro Thr Pro Glu Ala Asp Ala
 -165 -160 -155

Val Ser Met Gln Glu Ala Leu Gln Arg Asp Leu Gly Leu Thr Pro
 -150 -145 -140

Leu Glu Ala Asp Glu Leu Leu Ala Ala Gln Asp Thr Ala Phe Glu
 -135 -130 -125

Val Asp Glu Ala Ala Ala Ala Ala Ala Gly Asp Ala Tyr Gly Gly
 -120 -115 -110

Ser Val Phe Asp Thr Glu Thr Leu Glu Leu Thr Val Leu Val Thr Asp
 -105 -100 -95 -90

Ala Ala Ser Val Glu Ala Val Glu Ala Thr Gly Ala Gly Thr Glu Leu
 -85 -80 -75

Val Ser Tyr Gly Ile Glu Gly Leu Asp Glu Ile Ile Gln Asp Leu Asn
 -70 -65 -60

Ala Ala Asp Ala Val Pro Gly Val Val Gly Trp Tyr Pro Asp Val Ala
 -55 -50 -45

Gly Asp Thr Val Val Leu Glu Val Leu Glu Gly Ser Gly Ala Asp Val
 -40 -35 -30

Ser Gly Leu Leu Ala Asp Ala Gly Val Asp Ala Ser Ala Val Glu Val
 -25 -20 -15 -10

10423.204-WO.ST25.txt

Thr Ser Ser Ala Gln Pro Glu Leu Tyr Ala Asp Ile Ile Gly Gly Leu
 -5 -1 1 5

Ala Tyr Thr Met Gly Gly Arg Cys Ser Val Gly Phe Ala Ala Thr Asn
 10 15 20

Ala Ala Gly Gln Pro Gly Phe Val Thr Ala Gly His Cys Gly Arg Val
 25 30 35

Gly Thr Gln Val Ser Ile Gly Asn Gly Gln Gly Val Phe Glu Gln Ser
 40 45 50 55

Ile Phe Pro Gly Asn Asp Ala Ala Phe Val Arg Gly Thr Ser Asn Phe
 60 65 70

Thr Leu Thr Asn Leu Val Ser Arg Tyr Asn Thr Gly Gly Tyr Ala Thr
 75 80 85

Val Ala Gly His Asn Gln Ala Pro Ile Gly Ser Ser Val Cys Arg Ser
 90 95 100

Gly Ser Thr Thr Gly Trp His Cys Gly Thr Ile Gln Ala Arg Gly Gln
 105 110 115

Ser Val Ser Tyr Pro Glu Gly Thr Val Thr Asn Met Thr Arg Thr Thr
 120 125 130 135

Val Cys Ala Glu Pro Gly Asp Ser Gly Gly Ser Tyr Ile Ser Gly Asn
 140 145 150

Gln Ala Gln Gly Val Thr Ser Gly Gly Ser Gly Asn Cys Arg Thr Gly
 155 160 165

Gly Thr Thr Phe Tyr Gln Glu Val Thr Pro Met Val Asn Ser Trp Gly
 170 175 180

Val Arg Leu Arg Thr
 185

<210> 38

<211> 43

<212> DNA

<213> Artificial sequence

<220>

<223> Primer 1346

<400> 38

gttcatcgat cgcatcggtt gccaccggac cgctccccca gtc

43

<210> 39

<211> 38

10423.204-WO.ST25.txt

<212> DNA
 <213> Artificial sequence

<220>
 <223> Primer 1602

<400> 39
 gcggatccta ttaggtccgg agacggacgc cccaggag

38

<210> 40
 <211> 1062
 <212> DNA
 <213> Nocardiopsis prasina DSM 15649

<400> 40
 gccaccggac cactccccca gtcacccacc ccggaggccg acgccgtctc catgcaggag 60
 gcgctccagc gcgacctcgg cctgaccccg cttgaggccg atgaactgt ggccgcccag 120
 gacaccgcct tcgaggtcga cgaggccgcg gccgaggccg ccggtgacgc ctacggcggc 180
 tccgtttcg acaccgagac ccttggaaactg accgtcctgg tcaccgactc cgccgcggtc 240
 gaggcggtgg aggccaccgg cgccgggacc gaactggtct cctacggcat cacgggcctc 300
 gacgagatcg tcgaggagct caacgcccgc gacgcccgttc ccggcgtggt cggctggtac 360
 ccggacgtcg cgggtgacac cgtcgtgctg gaggtcctgg agggttccgg cgccgacgtg 420
 ggcggcctgc tcgcccacgc cggcgtggac gcctcggcgg tcgaggtgac caccaccgag 480
 cagcccggc tgtacgcccga catcatcgcc ggtctggcct acaccatggg cggccgtgt 540
 tcggtcggct tcgcccac caacgcccgc ggtcagcccg ggttcgtcac cgccggtcac 600
 tgtggccgcg tgggcaccca ggtgaccatc ggcaacggcc ggggcgtctt cgagcagtcc 660
 atcttcccgg gcaacgacgc cgccttcgtc cgcggAACgt ccaacttcac gctgaccaac 720
 ctggtcagcc gctacaacac cggcggctac gccaccgtcg ccggtcacaa ccaggcggccc 780
 atcggctcct ccgtctgccc ctccggctcc accaccgggtt ggcactgcgg caccatccag 840
 gcccggcggcc agtcggtgag ctaccccgag ggcaccgtca ccaacatgac gcggaccacc 900
 gtgtgcggcg agcccggcga ctccggcggc tcctacatct ccggcaacca ggcccagggc 960
 gtcacctccg cggcgtccgg caactgcccgc accggcggga ccaccttcta ccaggaggtc 1020
 accccatgg tgaactcctg gggcgtccgt ctccggacct aa 1062

<210> 41
 <211> 353
 <212> PRT
 <213> Nocardiopsis prasina DSM 15649

<220>
 <221> PROPEP
 <222> (1)..(165)

<220>
 <221> mat_peptide
 <222> (166)..(353)

<400> 41

10423.204-WO.ST25.txt

Ala Thr Gly Pro Leu Pro Gln Ser Pro Thr Pro Glu Ala Asp Ala
 -165 -160 -155

Val Ser Met Gln Glu Ala Leu Gln Arg Asp Leu Gly Leu Thr Pro
 -150 -145 -140

Leu Glu Ala Asp Glu Leu Leu Ala Ala Gln Asp Thr Ala Phe Glu
 -135 -130 -125

Val Asp Glu Ala Ala Ala Glu Ala Ala Gly Asp Ala Tyr Gly Gly
 -120 -115 -110

Ser Val Phe Asp Thr Glu Thr Leu Glu Leu Thr Val Leu Val Thr Asp
 -105 -100 -95 -90

Ser Ala Ala Val Glu Ala Val Glu Ala Thr Gly Ala Gly Thr Glu Leu
 -85 -80 -75

Val Ser Tyr Gly Ile Thr Gly Leu Asp Glu Ile Val Glu Glu Leu Asn
 -70 -65 -60

Ala Ala Asp Ala Val Pro Gly Val Val Gly Trp Tyr Pro Asp Val Ala
 -55 -50 -45

Gly Asp Thr Val Val Leu Glu Val Leu Glu Gly Ser Gly Ala Asp Val
 -40 -35 -30

Gly Gly Leu Leu Ala Asp Ala Gly Val Asp Ala Ser Ala Val Glu Val
 -25 -20 -15 -10

Thr Thr Thr Glu Gln Pro Glu Leu Tyr Ala Asp Ile Ile Gly Gly Leu
 -5 -1 1 5

Ala Tyr Thr Met Gly Gly Arg Cys Ser Val Gly Phe Ala Ala Thr Asn
 10 15 20

Ala Ala Gly Gln Pro Gly Phe Val Thr Ala Gly His Cys Gly Arg Val
 25 30 35

Gly Thr Gln Val Thr Ile Gly Asn Gly Arg Gly Val Phe Glu Gln Ser
 40 45 50 55

Ile Phe Pro Gly Asn Asp Ala Ala Phe Val Arg Gly Thr Ser Asn Phe
 60 65 70

Thr Leu Thr Asn Leu Val Ser Arg Tyr Asn Thr Gly Gly Tyr Ala Thr
 75 80 85

Val Ala Gly His Asn Gln Ala Pro Ile Gly Ser Ser Val Cys Arg Ser
 90 95 100

10423.204-WO.ST25.txt

Gly Ser Thr Thr Gly Trp His Cys Gly Thr Ile Gln Ala Arg Gly Gln
 105 110 115

Ser Val Ser Tyr Pro Glu Gly Thr Val Thr Asn Met Thr Arg Thr Thr
 120 125 130 135

Val Cys Ala Glu Pro Gly Asp Ser Gly Gly Ser Tyr Ile Ser Gly Asn
 140 145 150

Gln Ala Gln Gly Val Thr Ser Gly Gly Ser Gly Asn Cys Arg Thr Gly
 155 160 165

Gly Thr Thr Phe Tyr Gln Glu Val Thr Pro Met Val Asn Ser Trp Gly
 170 175 180

Val Arg Leu Arg Thr
 185

<210> 42
 <211> 43
 <212> DNA
 <213> Artificial sequence

<220>
 <223> Primer 1603

<400> 42
 gttcatcgat cgcatcggt gccaccggac cactcccca gtc

43

<210> 43
 <211> 353
 <212> PRT
 <213> Nocardiopsis sp. NRRL 18262

<220>
 <221> PROPEP
 <222> (1)..(165)

<220>
 <221> mat_peptide
 <222> (166)..(1059)

<400> 43

Ala Thr Gly Ala Leu Pro Gln Ser Pro Thr Pro Glu Ala Asp Ala
 -165 -160 -155

Val Ser Met Gln Glu Ala Leu Gln Arg Asp Leu Asp Leu Thr Ser
 -150 -145 -140

Ala Glu Ala Glu Glu Leu Leu Ala Ala Gln Asp Thr Ala Phe Glu
 -135 -130 -125

Val Asp Glu Ala Ala Ala Glu Ala Ala Gly Asp Ala Tyr Gly Gly
 -120 -115 -110

10423.204-WO.ST25.txt

Ser Val Phe Asp Thr Glu Ser Leu Glu Leu Thr Val Leu Val Thr Asp
 -105 -100 -95 -90

Ala Ala Ala Val Glu Ala Val Glu Ala Thr Gly Ala Gly Thr Glu Leu
 -85 -80 -75

Val Ser Tyr Gly Ile Asp Gly Leu Asp Glu Ile Val Gln Glu Leu Asn
 -70 -65 -60

Ala Ala Asp Ala Val Pro Gly Val Val Gly Trp Tyr Pro Asp Val Ala
 -55 -50 -45

Gly Asp Thr Val Val Leu Glu Val Leu Glu Gly Ser Gly Ala Asp Val
 -40 -35 -30

Ser Gly Leu Leu Ala Asp Ala Gly Val Asp Ala Ser Ala Val Glu Val
 -25 -20 -15 -10

Thr Thr Ser Asp Gln Pro Glu Leu Tyr Ala Asp Ile Ile Gly Gly Leu
 -5 -1 1 5

Ala Tyr Thr Met Gly Gly Arg Cys Ser Val Gly Phe Ala Ala Thr Asn
 10 15 20

Ala Ala Gly Gln Pro Gly Phe Val Thr Ala Gly His Cys Gly Arg Val
 25 30 35

Gly Thr Gln Val Thr Ile Gly Asn Gly Arg Gly Val Phe Glu Gln Ser
 40 45 50 55

Val Phe Pro Gly Asn Asp Ala Ala Phe Val Arg Gly Thr Ser Asn Phe
 60 65 70

Thr Leu Thr Asn Leu Val Ser Arg Tyr Asn Thr Gly Gly Tyr Ala Thr
 75 80 85

Val Ala Gly His Asn Gln Ala Pro Ile Gly Ser Ser Val Cys Arg Ser
 90 95 100

Gly Ser Thr Thr Gly Trp His Cys Gly Thr Ile Gln Ala Arg Gly Gln
 105 110 115

Ser Val Ser Tyr Pro Glu Gly Thr Val Thr Asn Met Thr Arg Thr Thr
 120 125 130 135

Val Cys Ala Glu Pro Gly Asp Ser Gly Gly Ser Tyr Ile Ser Gly Thr
 140 145 150

Gln Ala Gln Gly Val Thr Ser Gly Gly Ser Gly Asn Cys Arg Thr Gly
 155 160 165

10423.204-WO.ST25.txt

Gly Thr Thr Phe Tyr Gln Glu Val Thr Pro Met Val Asn Ser Trp Gly
170 175 180

Val Arg Leu Arg Thr
185

<210> 44
<211> 1164
<212> DNA
<213> artificial sequence

<220>
<223> Synthetic protease encoding gene

<220>
<221> CDS
<222> (1)..(1164)
<223> Full length protease

<220>
<221> sig_peptide
<222> (1)..(81)

<220>
<221> misc_feature
<222> (82)..(1164)
<223> Propeptide

<220>
<221> mat_peptide
<222> (577)..(1164)

<400> 44
atg aaa aaa ccg ctg gga aaa att gtc gca agc aca gca ctt ctt
Met Lys Lys Pro Leu Gly Lys Ile Val Ala Ser Thr Ala Leu Leu
-190 -185 -180 45

att tca gtg gca ttt agc tca tct att gca tca gca gct aca gga
 Ile Ser Val Ala Phe Ser Ser Ser Ile Ala Ser Ala Ala Thr Gly
 -175 -170 -165 90

gca tta ccg cag tct ccg aca ccg gaa gca gat gca gtc tca atg
 Ala Leu Pro Gln Ser Pro Thr Pro Glu Ala Asp Ala Val Ser Met
 -160 -155 -150 135

caa gaa gca ctg caa aga gat ctt gat ctt aca tca gca gaa gca
 Gln Glu Ala Leu Gln Arg Asp Leu Asp Leu Thr Ser Ala Glu Ala
 -145 -140 -135 180

gaa gaa ctt ctt gct gca caa gat aca gca ttt gaa gtg gat gaa
 Glu Glu Leu Leu Ala Ala Gln Asp Thr Ala Phe Glu Val Asp Glu 225
 -130 -125 -120

gca gcg gca gaa gca gca gya gat gca tat ggc ggc tca gtt ttt
 Ala Ala Ala Glu Ala Ala Gly Asp Ala Tyr Gly Gly Ser Val Phe
 -115 -110 -105
 270

gat aca gaa tca ctt gaa ctt aca gtt ctt gtt aca gat gca gca gca
Asp Thr Glu Ser Leu Glu Leu Thr Val Leu Val Thr Asp Ala Ala Ala
-100 -95 -90

gtt gaa gca gtt gaa gca aca gga gca gga aca gta ctt gtt tca tat	366
Val Glu Ala Val Glu Ala Thr Gly Ala Gly Thr Val Leu Val Ser Tyr	
-85 -80 -75	

10423.204-WO.ST25.txt

gga att gat ggc ctt gat gaa att gtt caa gaa ctg aat gca gct gat	Gly Ile Asp Gly Leu Asp Glu Ile Val Gln Glu Leu Asn Ala Ala Asp	414
-70 -65 -60 -55		
gct gtt ccg ggc gtt gtc tat ccg gat gtt gct gga gat aca	Ala Val Pro Gly Val Val Gly Trp Tyr Pro Asp Val Ala Gly Asp Thr	462
-50 -45 -40		
gtt gtc ctt gaa gtt ctt gaa gga tca ggc gca gat gtt tca ggc ctg	Val Val Leu Glu Val Leu Gly Ser Gly Ala Asp Val Ser Gly Leu	510
-35 -30 -25		
ctg gca gac gca gga gtc gat gca tca gca gtt gaa gtt aca aca tca	Leu Ala Asp Ala Gly Val Asp Ala Ser Ala Val Glu Val Thr Thr Ser	558
-20 -15 -10		
gat caa ccg gaa ctt tat gca gat att att ggc ggc ctg gca tat tat	Asp Gln Pro Glu Leu Tyr Ala Asp Ile Ile Gly Gly Leu Ala Tyr Tyr	606
-5 -1 1 5 10		
atg ggc ggc aga tgc agc gtt ggc ttt gca gca aca aat gca tca ggc	Met Gly Gly Arg Cys Ser Val Gly Phe Ala Ala Thr Asn Ala Ser Gly	654
15 20 25		
caa ccg ggc ttt gtt aca gca ggc cat tgc ggc aca gtt ggc aca cca	Gln Pro Gly Phe Val Thr Ala Gly His Cys Gly Thr Val Gly Thr Pro	702
30 35 40		
gtt tca att ggc aat ggc aaa ggc gtt ttt gaa cga agc att ttt ccg	Val Ser Ile Gly Asn Gly Lys Gly Val Phe Glu Arg Ser Ile Phe Pro	750
45 50 55		
ggc aat gat tca gca ttt gtt aga ggc aca tca aat ttt aca ctt aca	Gly Asn Asp Ser Ala Phe Val Arg Gly Thr Ser Asn Phe Thr Leu Thr	798
60 65 70		
aat ctg gtt tca aga tat aat tca ggc ggc tat gca aca gtt gca ggc	Asn Leu Val Ser Arg Tyr Asn Ser Gly Gly Tyr Ala Thr Val Ala Gly	846
75 80 85 90		
cat aat caa gca ccg att ggc tca gca gtt tgc aga tca ggc tca aca	His Asn Gln Ala Pro Ile Gly Ser Ala Val Cys Arg Ser Gly Ser Thr	894
95 100 105		
aca ggc tgg cat tgc ggc aca att caa gca aga aat caa aca gtt agg	Thr Gly Trp His Cys Gly Thr Ile Gln Ala Arg Asn Gln Thr Val Arg	942
110 115 120		
tat ccg caa ggc aca gtt tat agt ctg aca aga aca aca gtt tgt gca	Tyr Pro Gln Gly Thr Val Tyr Ser Leu Thr Arg Thr Thr Val Cys Ala	990
125 130 135		
gaa ccg ggc gat tca ggc tca tat att agc ggc act caa gca caa	Glu Pro Gly Asp Ser Gly Gly Ser Tyr Ile Ser Gly Thr Gln Ala Gln	1038
140 145 150		
ggc gtt aca tca ggc ggc tca ggc aat tgc agt gct ggc ggc aca aca	Gly Val Thr Ser Gly Gly Ser Gly Asn Cys Ser Ala Gly Gly Thr Thr	1086
155 160 165 170		
tat tac caa gaa gtt aat ccg atg ctt agt tca tgg ggc ctt aca ctt	Tyr Tyr Gln Glu Val Asn Pro Met Leu Ser Ser Trp Gly Leu Thr Leu	1134
175 180 185		
aga aca caa tcg cat gtt caa tcc gct cca	Arg Thr Gln Ser His Val Gln Ser Ala Pro	1164
190 195		

10423.204-WO.ST25.txt

<210> 45
 <211> 388
 <212> PRT
 <213> artificial sequence

<220>
 <223> Synthetic Construct
 <400> 45

Met Lys Lys Pro Leu Gly Lys Ile Val Ala Ser Thr Ala Leu Leu
 -190 -185 -180

Ile Ser Val Ala Phe Ser Ser Ser Ile Ala Ser Ala Ala Thr Gly
 -175 -170 -165

Ala Leu Pro Gln Ser Pro Thr Pro Glu Ala Asp Ala Val Ser Met
 -160 -155 -150

Gln Glu Ala Leu Gln Arg Asp Leu Asp Leu Thr Ser Ala Glu Ala
 -145 -140 -135

Glu Glu Leu Leu Ala Ala Gln Asp Thr Ala Phe Glu Val Asp Glu
 -130 -125 -120

Ala Ala Ala Glu Ala Ala Gly Asp Ala Tyr Gly Gly Ser Val Phe
 -115 -110 -105

Asp Thr Glu Ser Leu Glu Leu Thr Val Leu Val Thr Asp Ala Ala Ala
 -100 -95 -90

Val Glu Ala Val Glu Ala Thr Gly Ala Gly Thr Val Leu Val Ser Tyr
 -85 -80 -75

Gly Ile Asp Gly Leu Asp Glu Ile Val Gln Glu Leu Asn Ala Ala Asp
 -70 -65 -60 -55

Ala Val Pro Gly Val Val Gly Trp Tyr Pro Asp Val Ala Gly Asp Thr
 -50 -45 -40

Val Val Leu Glu Val Leu Glu Gly Ser Gly Ala Asp Val Ser Gly Leu
 -35 -30 -25

Leu Ala Asp Ala Gly Val Asp Ala Ser Ala Val Glu Val Thr Thr Ser
 -20 -15 -10

Asp Gln Pro Glu Leu Tyr Ala Asp Ile Ile Gly Gly Leu Ala Tyr Tyr
 -5 -1 1 5 10

Met Gly Gly Arg Cys Ser Val Gly Phe Ala Ala Thr Asn Ala Ser Gly
 15 20 25

10423.204-WO.ST25.txt
 Gln Pro Gly Phe Val Thr Ala Gly His Cys Gly Thr Val Gly Thr Pro
 30 35 40

Val Ser Ile Gly Asn Gly Lys Gly Val Phe Glu Arg Ser Ile Phe Pro
 45 50 55

Gly Asn Asp Ser Ala Phe Val Arg Gly Thr Ser Asn Phe Thr Leu Thr
 60 65 70

Asn Leu Val Ser Arg Tyr Asn Ser Gly Gly Tyr Ala Thr Val Ala Gly
 75 80 85 90

His Asn Gln Ala Pro Ile Gly Ser Ala Val Cys Arg Ser Gly Ser Thr
 95 100 105

Thr Gly Trp His Cys Gly Thr Ile Gln Ala Arg Asn Gln Thr Val Arg
 110 115 120

Tyr Pro Gln Gly Thr Val Tyr Ser Leu Thr Arg Thr Thr Val Cys Ala
 125 130 135

Glu Pro Gly Asp Ser Gly Gly Ser Tyr Ile Ser Gly Thr Gln Ala Gln
 140 145 150

Gly Val Thr Ser Gly Gly Ser Gly Asn Cys Ser Ala Gly Gly Thr Thr
 155 160 165 170

Tyr Tyr Gln Glu Val Asn Pro Met Leu Ser Ser Trp Gly Leu Thr Leu
 175 180 185

Arg Thr Gln Ser His Val Gln Ser Ala Pro
 190 195

<210> 46
 <211> 165
 <212> PRT
 <213> Artificial sequence

<220>
 <223> Shuffled pro-peptide 0-2.19

<220>
 <221> PROPEP
 <222> (1)..(165)

<400> 46

Ala Thr Gly Ala Leu Pro Gln Ser Pro Thr Pro Glu Ala Asp Ala Val
 1 5 10 15

Ser Met Gln Glu Ala Leu Gln Arg Asp Leu Asp Leu Thr Ser Ala Glu
 20 25 30

Ala Glu Glu Leu Leu Ala Ala Gln Asp Thr Ala Phe Glu Val Asp Glu
 Page 30

35

10423.204-WO.ST25.txt
40 45

Ala Ala Ala Ala Ala Ala Gly Asp Ala Tyr Gly Gly Ser Val Phe Asp
50 55 60

Thr Glu Ser Leu Thr Leu Thr Val Leu Val Thr Asp Ala Ser Ala Val
65 70 75 80

Glu Ala Val Glu Ala Ala Gly Ala Glu Ala Lys Val Val Ser His Gly
85 90 95

Met Glu Gly Leu Glu Glu Ile Val Ala Asp Leu Asn Ala Ala Asp Ala
100 105 110

Gln Pro Gly Val Val Gly Trp Tyr Pro Asp Ile His Ser Asp Thr Val
115 120 125

Val Leu Glu Val Leu Glu Gly Ser Gly Ala Asp Val Asp Ser Leu Leu
130 135 140

Ala Asp Ala Gly Val Asp Ala Ser Ala Val Glu Val Thr Thr Ser Asp
145 150 155 160

Gln Pro Glu Leu Tyr
165

<210> 47

<211> 166

<212> PRT

<213> Artificial sequence

<220>

<223> Shuffled propeptide G-2.73

<220>

<221> PROPEP

<222> (1)..(166)

<400> 47

Ala Thr Gly Ala Leu Pro Gln Ser Pro Thr Pro Glu Ala Asp Ala Val
1 5 10 15

Ser Met Gln Glu Ala Leu Gln Arg Asp Leu Asp Leu Ser Ser Ala Glu
20 25 30

Ala Glu Glu Leu Leu Ala Ala Gln Asp Thr Ala Phe Glu Val Asp Glu
35 40 45

Ala Ala Ala Gly Ala Ala Gly Asp Ala Tyr Gly Gly Ser Val Phe Asp
50 55 60

Thr Glu Thr Leu Glu Leu Thr Val Leu Val Thr Asp Ala Ser Ala Val
65 70 75 80

10423.204-WO.ST25.txt

Glu Ala Val Glu Ala Ala Gly Ala Glu Ala Lys Val Val Ser His Gly
85 90 95

Met Glu Gly Leu Glu Glu Ile Val Ala Asp Leu Asn Ala Ala Asp Ala
100 105 110

Gln Pro Gly Val Val Gly Trp Tyr Pro Asp Ile His Ser Asp Thr Val
115 120 125

Val Val Glu Val Leu Glu Gly Ser Gly Ala Asp Val Asp Ser Leu Leu
130 135 140

Ala Asp Ala Gly Val Asp Thr Ala Asp Val Lys Val Glu Ser Thr Thr
145 150 155 160

Glu Gln Pro Glu Leu Tyr
165

<210> 48

<211> 166

<212> PRT

<213> Artificial sequence

<220>

<223> Shuffled propeptide G-1.43

<220>

<221> PROPEP

<222> (1)..(166)

<400> 48

Ala Thr Gly Ala Leu Pro Gln Ser Pro Thr Pro Glu Ala Asp Ala Val
1 5 10 15

Ser Met Gln Glu Ala Leu Gln Arg Asp Leu Gly Leu Ser Ser Ser Gln
20 25 30

Ala Glu Glu Leu Leu Asp Ala Gln Ala Glu Ser Phe Glu Ile Asp Glu
35 40 45

Ala Ala Ala Ala Ala Ala Gly Asp Ala Tyr Gly Ser Ile Phe Asp
50 55 60

Thr Asp Ser Leu Thr Leu Thr Val Leu Val Thr Asp Ala Ser Ala Val
65 70 75 80

Glu Ala Val Glu Ala Ala Gly Ala Glu Ala Lys Val Val Ser His Gly
85 90 95

Met Glu Gly Leu Glu Glu Ile Val Ala Asp Leu Asn Ala Ala Asp Ala
100 105 110

10423.204-WO.ST25.txt

Gln Pro Gly Val Val Gly Trp Tyr Pro Asp Ile His Ser Asp Thr Val
 115 120 125

Val Leu Glu Val Leu Glu Gly Ser Gly Ala Asp Val Asp Ser Leu Leu
 130 135 140

Ala Asp Ala Gly Val Asp Thr Ala Asp Val Lys Val Glu Ser Thr Thr
 145 150 155 160

Glu Gln Pro Glu Leu Tyr
 165

<210> 49

<211> 166

<212> PRT

<213> Artificial sequence

<220>

<223> shuffled propeptide G-2.6

<220>

<221> PROPEP

<222> (1)..(166)

<400> 49

Ala Thr Gly Ala Leu Pro Gln Ser Pro Thr Pro Glu Ala Asp Ala Val
 1 5 10 15

Ser Met Gln Glu Ala Leu Gln Arg Asp Leu Asp Leu Thr Ser Ala Glu
 20 25 30

Ala Glu Glu Leu Leu Ala Ala Gln Asp Thr Ala Phe Glu Val Asp Glu
 35 40 45

Ala Ala Ala Ala Ala Ala Gly Asp Ala Tyr Gly Gly Ser Ile Phe Asp
 50 55 60

Thr Glu Thr Leu Glu Leu Thr Val Leu Val Thr Asp Ser Ser Ser Val
 65 70 75 80

Glu Ala Val Glu Ala Ala Gly Ala Glu Ala Lys Val Val Ser His Gly
 85 90 95

Met Glu Gly Leu Glu Glu Ile Val Ala Asp Leu Asn Ala Ala Asp Ala
 100 105 110

Gln Pro Gly Val Val Gly Trp Tyr Pro Asp Ile His Ser Asp Thr Val
 115 120 125

Val Leu Glu Val Leu Glu Gly Ser Gly Ala Asp Val Asp Ser Leu Leu
 130 135 140

10423.204-WO.ST25.txt
Ala Gly Ala Gly Val Asp Thr Ala Asp Val Lys Val Glu Ser Thr Thr
145 150 155 160

Glu Gln Pro Glu Leu Tyr
165

<210> 50
<211> 165
<212> PRT
<213> Artificial sequence

<220>
<223> Shuffled propeptide G-2.5

<220>
<221> PROPEP
<222> (1)..(165)

<400> 50

Ala Thr Gly Ala Leu Pro Gln Ser Pro Thr Pro Glu Ala Asp Ala Val
1 5 10 15

Ser Met Gln Glu Ala Leu Gln Arg Asp Leu Gly Leu Thr Pro Leu Glu
20 25 30

Ala Glu Glu Leu Leu Ala Ala Gln Asp Thr Ala Phe Glu Val Asp Glu
35 40 45

Ala Ala Ala Glu Ala Ala Gly Asp Ala Tyr Gly Ser Val Phe Asp
50 55 60

Thr Glu Thr Leu Glu Leu Thr Val Leu Val Thr Asp Ala Ser Ala Val
65 70 75 80

Glu Ala Val Glu Ala Ala Gly Ala Glu Ala Lys Val Val Ser His Gly
85 90 95

Met Glu Gly Leu Glu Glu Ile Val Ala Asp Leu Asn Ala Ala Asp Ala
100 105 110

Gln Pro Gly Val Val Gly Trp Tyr Pro Asp Ile His Ser Asp Thr Val
115 120 125

Val Leu Glu Val Leu Glu Gly Ser Gly Ala Asp Val Asp Ser Leu Leu
130 135 140

Ala Asp Ala Gly Val Asp Ala Ser Ala Val Glu Val Thr Pro Ala Ala
145 150 155 160

Arg Pro Glu Leu Tyr
165

<210> 51

10423.204-WO.ST25.txt

<211> 166
 <212> PRT
 <213> Artificial sequence

<220>
 <223> Shuffled propeptide G-2.3

<220>
 <221> PROPEP
 <222> (1)..(166)

<400> 51

Ala Thr Gly Ala Leu Pro Gln Ser Pro Thr Pro Asp Gly Ala Glu Ala
 1 5 10 15

Thr Thr Met Val Glu Ala Leu Gln Arg Asp Leu Gly Leu Thr Pro Ala
 20 25 30

Glu Ala Glu Glu Leu Leu Ala Ala Gln Asp Thr Ala Phe Glu Val Asp
 35 40 45

Glu Ala Ala Ala Ala Ala Gly Asp Ala Tyr Gly Gly Ser Ile Phe
 50 55 60

Asp Thr Asp Ser Leu Thr Leu Thr Val Leu Val Thr Asp Ala Ala Ala
 65 70 75 80

Val Glu Ala Val Glu Ala Ala Gly Ala Glu Ala Lys Val Val Ser His
 85 90 95

Gly Met Glu Gly Leu Glu Glu Ile Val Ala Asp Leu Asn Ala Ala Asp
 100 105 110

Ala Val Pro Gly Val Val Gly Trp Tyr Pro Asp Val Ala Gly Asp Thr
 115 120 125

Val Val Leu Glu Val Leu Glu Gly Ser Gly Ala Asp Val Tyr Ser Leu
 130 135 140

Leu Ala Asp Ala Gly Val Asp Ala Ser Ala Val Glu Val Thr Pro Ala
 145 150 155 160

Ala Gln Pro Glu Leu Tyr
 165

<210> 52
 <211> 166
 <212> PRT
 <213> Artificial sequence

<220>
 <223> Shuffled propeptide G-1.4

<220>

10423.204-WO.ST25.txt

<221> PROPEP
<222> (1)..(166)

<400> 52

Ala Thr Gly Ala Leu Pro Gln Ser Pro Thr Pro Glu Ala Asp Ala Val
1 5 10 15

Ser Met Gln Glu Ala Leu Gln Arg Asp Leu Gly Leu Ser Ser Ser Gln
20 25 30

Ala Glu Glu Leu Leu Asp Ala Gln Ala Glu Ser Phe Glu Ile Asp Glu
35 40 45

Ala Ala Ala Ala Ala Ala Asp Ser Tyr Gly Gly Ser Ile Phe Asp
50 55 60

Thr Asp Ser Leu Thr Leu Thr Val Leu Val Thr Asp Ala Ser Ala Val
65 70 75 80

Glu Ala Val Glu Ala Ala Gly Ala Glu Ala Lys Val Val Ser His Gly
85 90 95

Met Glu Gly Leu Glu Glu Ile Val Ala Asp Leu Asn Ala Ala Asp Ala
100 105 110

Gln Pro Gly Val Val Gly Trp Tyr Pro Asp Ile His Ser Asp Thr Val
115 120 125

Val Leu Glu Val Leu Glu Gly Ser Gly Ala Asp Val Asp Ser Leu Leu
130 135 140

Ala Asp Ala Gly Val Asp Thr Ala Asp Val Lys Val Glu Ser Thr Thr
145 150 155 160

Glu Gln Pro Glu Leu Tyr
165

<210> 53
<211> 166
<212> PRT
<213> Artificial sequence

<220>
<223> Shuffled propeptide G-1.2

<220>
<221> PROPEP
<222> (1)..(166)

<400> 53

Ala Thr Gly Ala Leu Pro Gln Ser Pro Thr Pro Glu Ala Asp Ala Val
1 5 10 15

10423.204-WO.ST25.txt

Ser Met Gln Glu Ala Leu Gln Arg Asp Leu Asp Leu Thr Ser Ala Glu
20 25 30

Ala Glu Glu Leu Leu Ala Ala Gln Asp Thr Ala Phe Glu Val Asp Glu
35 40 45

Ala Ala Ala Ala Ala Ala Gly Asp Ala Tyr Gly Gly Ser Ile Phe Asp
50 55 60

Thr Glu Thr Leu Glu Leu Thr Val Leu Val Thr Asp Ser Ser Ser Val
65 70 75 80

Glu Ala Val Glu Ala Ala Gly Ala Glu Ala Lys Val Val Ser His Gly
85 90 95

Met Glu Gly Leu Glu Glu Ile Val Ala Asp Leu Asn Ala Ala Asp Ala
100 105 110

Gln Pro Gly Val Val Gly Trp Tyr Pro Asp Ile His Ser Asp Thr Val
115 120 125

Val Leu Glu Val Leu Glu Gly Ser Gly Ala Asp Val Asp Ser Leu Leu
130 135 140

Ala Gly Ala Gly Val Asp Thr Ala Asp Val Lys Val Glu Ser Thr Thr
145 150 155 160

Glu Gln Pro Glu Leu Tyr
165